

# Wide Area Information Servers

Brewster Kahle  
Thinking Machines Corporation  
March 15, 1991

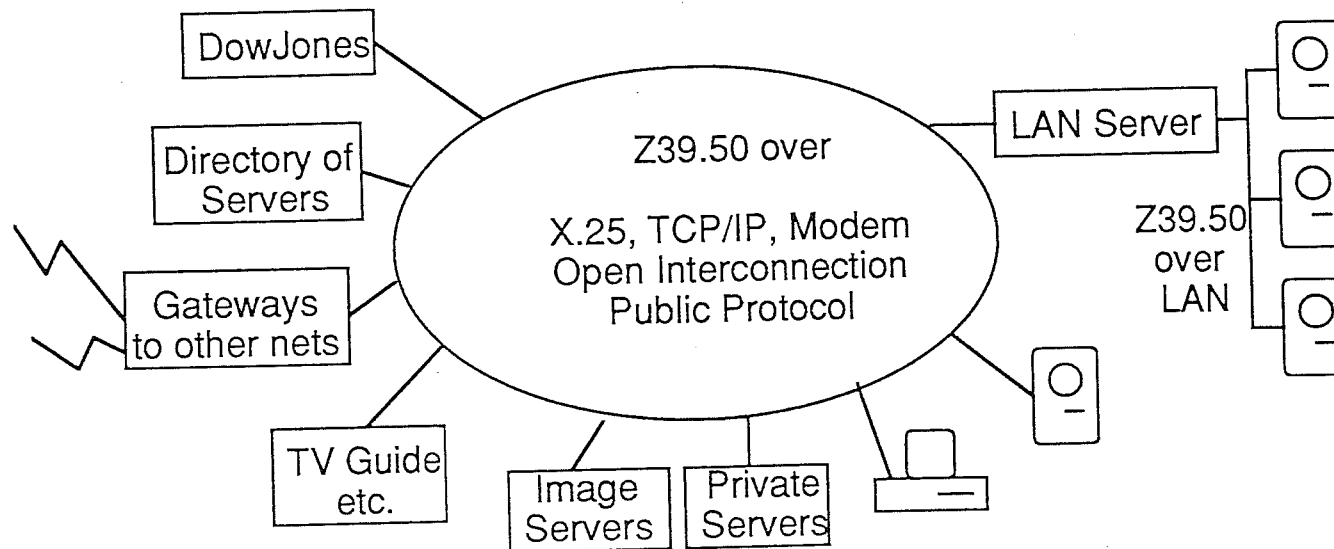
- Goals
- WAIS Architecture
- Proposed Techniques
- Demonstration System Architecture
- Project Structure
- Thinking Machines Plans
- Conclusion

## Goals, and where are we?

- **Catalyze a market of Wide Area Information Servers:**  
*status: This is a long road.*
- **Establish open protocol for Information Servers:**  
*status: done. Z39.50 is catching on. Our changes are being adopted*
- **Produce skeleton Producer ---> Consumer system:**  
*status: done.*
- **If the system is useful:**  
Move into production at TMC and Dow: *in progress*  
Transfer technology to Apple Products Division: *maybe*  
Evangelize third parties to write Mac software: *starting*
- **Find out what is missing from our system:**  
*status: lets pull it together*

WAIS

# Wide Area Information Server Architecture



**Users Problems:**  
Selecting Servers  
Answering Questions  
Organizing Responses

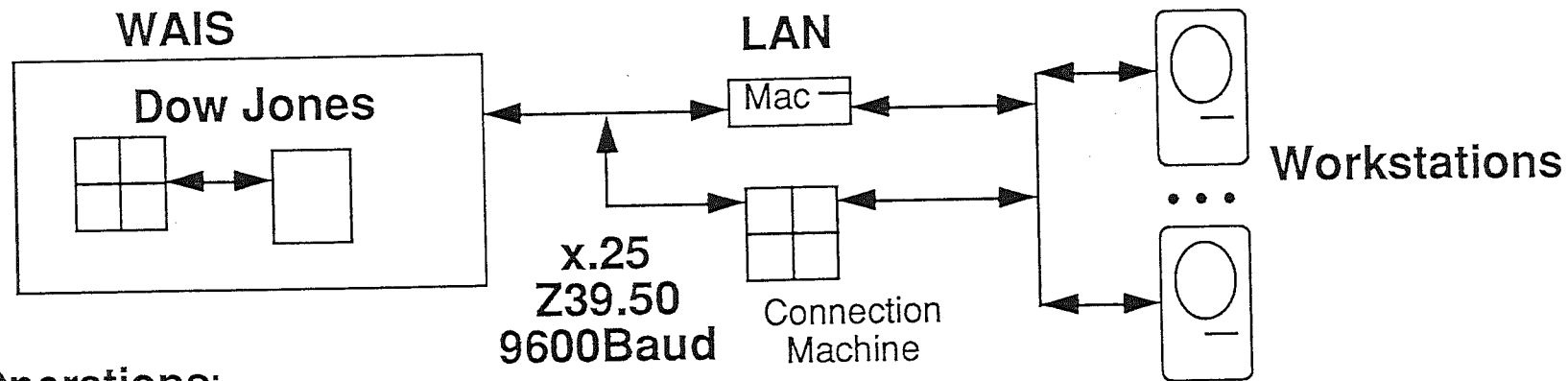
**Architecture Issues:**  
Scalability  
Security  
Business model for servers  
Reliable Access

## Proposed Techniques

- **English Questions yield "headlines" to guide retrieval; Query by example:**  
*status: Done. Needs refinement*
- **Dynamic folders hold responses and will work at night:** *status: never really worked.*
- **Directory of Servers can be questioned like any other server:** *status: just coming up now*
- **Copy Editors help select important information:**  
*status: not implemented.*
- **Single Computer-to-Computer Protocol:** *status: works*
- **Create small system to evaluate the design:** *status: done*

WAIS

# Demonstration System Structure



## Operations:

Archiving  
Queries  
Retrieval

## IR Type:

Broadcast  
Query by Example

## Databases:

Wall St Journal  
Barron's  
400 Business Mags

*\* italics means not done yet*

## Mac:

**Operations:** Archiving  
Queries  
Retrieval

**IR Type:** Query by Example

**DBs:** Mail, Netnews  
Encyclopedias  
DowVision

**CM: Operations:** Queries

## IR Type:

enhanced relevance feedback

**DBs:** DowVision and memo's

## Mac:

**Operations:** Human Int  
Retrieval  
Queries

"Caching" Docs  
User Profiles

## IR Type:

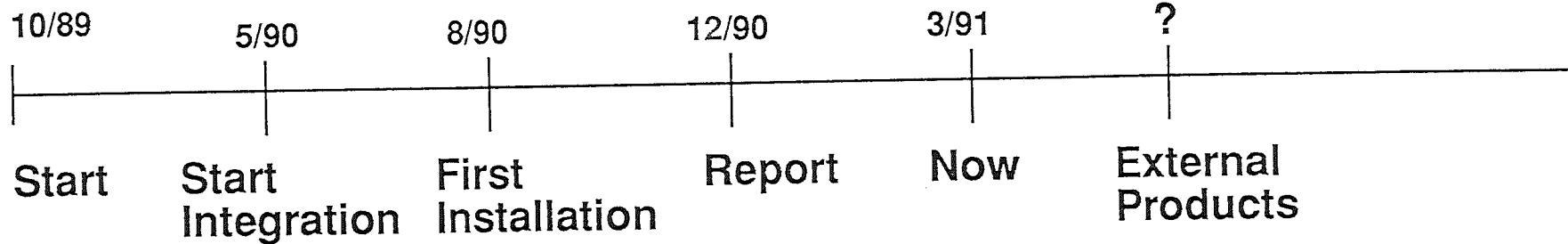
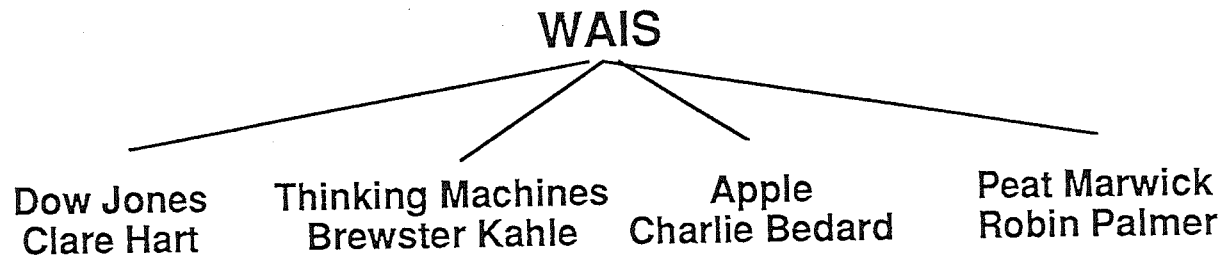
Query by example

## DBs:

Personal Text  
Cached data

WAIS

# Project Structure



## Thinking Machines Plans

- Index larger amounts of data
- Better retrieval on business information
- Support the protocol library at least through 1991
- Minimize effort on user interfaces
- Support a directory of servers through 1991

## Internet Release

- CM product for TCP/IP (complete server)
- Example User interfaces for free (no support)
  - Macintosh
  - Gnu Emacs
  - Xwindows
- Example unix server software to create servers:  
(no support)
- Directory of Servers on the internet at least through '91
- Alpha release 3/91, Beta 5/91



## Conclusion

- We succeeded beyond our expectations with our prototype: We have built what we set out to.
- The users and the technology seem ready for a set of products in this area.
- What are our plans for the future?